DOCUMENT RESUME

ED 356 052 PS 021 257

AUTHOR	Broadfoot, Patricia; And Others
TITLE	Look Back in Anger? Findings from the PACE Project Concerning Primary Teachers' Experiences of SATs.
INSTITUTION	Bristol Univ. (England). School of Education.
REPORT NO	PACE-WP-8
PUB DATE	91
NOTE	12p.
PUB TYPE	Reports - Research/Technical (143)
EDRS PRICE	MF01/PC01 Plus Postage.
DESCRIPTORS	Anxiety; *Elementary Education; Evaluation Methods; Evaluation Problems; Foreign Countries; *National Competency Tests; *Standardized Tests; *Student Evaluation; *Teacher Attitudes; *Teaching Experience
IDENTIFIERS	Primary Assessment Curric Exper (Great Britain); *Standard Assessment Tasks (Great Britain); Teacher Surveys; United Kingdom

ABSTRACT

A product of the United Kingdom's Primary Assessment, Curriculum and Experience (PACE) project, this report presents interview, observation, and survey findings concerning primary school teachers' attitudes toward and experiences of the Standard Assessment Tasks (SATs). In 1991, students were observed taking SATs at nine schools. Afterwards, students and teachers were interviewed about the experience. In addition, year 2 teachers in 48 schools completed questionnaires, and head teachers gave open-ended comments on their experiences. Teachers reported anxiety in several areas related to the test, including extra workloads, relationships with colleagues and parents, and disruption of normal teaching practices. Teachers also expressed concern about the reliability of SAT assessments. The report concludes by noting teachers' general willingness to accept SATs despite their negative experiences and criticisms. (MM)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED356052

PACE

PRIMARY ASSESSMENT, CURRICULUM AND EXPERIENCE

A study of educational change under the National Curriculum

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

X This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

SCOPE OF INTEREST NOTICE

The ERIC Facility has assigned this document for processing to: PS

In our judgment, this document is also of interest to the Clearinghouses noted to the right. Indexing should reflect their special points of view. TM

LOOK BACK IN ANGER?

FINDINGS FROM THE PACE PROJECT CONCERNING PRIMARY TEACHERS' EXPERIENCES OF SATs

by

Patricia Broadfoot and Dorothy Abbott
with Paul Croll, Marilyn Osborn and Andrew Pollard

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

M. Osborn

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

PACE Working Paper No 8

PS 021257

BEST COPY AVAILABLE

LOOK BACK IN ANGER?

Findings from the PACE project concerning primary teachers' experiences of SATs

Patricia Broadfoot and Dorothy Abbott
with Paul Croll, Marilyn Osborn and Andrew Pollard

For many primary school teachers their abiding memory of the summer term of 1991 will be SATs: learning about SATs; doing SATs; recording SATs, reporting SATs and coping with the consequences of SATs - for the children themselves; for relationships with parents and for the overall quality of school life. Some of these consequences have been good; many are clearly bad and no doubt more effects will become evident in the longer term.

Certainly there has been no shortage of feedback on teachers' experience of this year's SATs. The media have given extensive coverage of some of the more obvious problems of teacher workload and parent and pupil anxiety. But now that the dust is beginning to settle just a little and with the prospect of a rather different kind of SAT on the horizon, it is appropriate to consider in a rather more measured way what we may have learned from an experience that was undoubtedly cathartic for schools.

In the Primary Assessment, Curriculum and Experience (PACE) project, our aim is to provide comprehensive documentation of these effects through regular observation of a national sample of infant classrooms and discussions with teachers and pupils. During the first part of the summer term 1991 members of the PACE team visited nine schools where we observed individual children and groups carrying out SATs, and interviewed them and their teachers about the experience. In addition Year 2 teachers in all 48 schools in the full PACE sample have been asked to complete questionnaires while heads have been asked for open-ended comments on their experiences. This article is based on the results of these enquiries.

Firstly, it is important to stress the broad elements in how this substantial extra workload was made manageable by schools and teachers. 'The SAT is designed to be used by a single teacher with the resources normally available in schools' suggests the Handbook of Guidance for SAT, (p.12). Most

SATs were indeed based on activities frequently seen in infant classrooms and the resources needed were usually easily available. The great exception was time: although SATs had been widely piloted, most teachers found that tasks took far longer than the estimates in the handbook. To give individual children and small groups the close attention needed for careful assessment, much extra support and help were needed and unaided teachers usually found it very difficult to integrate the SAT group with the rest of the class.

The degree of difficulty varied with the task, of course; children engaged in making a game (Maths I, Part B) were able to spend some time without close supervision while drawing and colouring their games. Exploring floating and sinking, however (Science I, Part B) made heavy and virtually constant demands on the teacher's time and attention, quite possible when the SAT was carried out in a separate room or when the rest of the class was otherwise supervised but extremely difficult for teachers who tried to supervise the SAT and the rest of the class at the same time.

Every Year 2 teacher in the PACE survey was involved in carrying out SATs herself with about half being helped by the head or another colleague undertaking some of the assessments. Virtually all classrooms - 92% - had SATs going on in them though many schools were also able to use a separate place such as a library or staffroom. Only 19% of teachers had no other help in the classroom while the SATs were taking place. The vast majority had the support of a supply teacher, an assistant, a student, parents or a colleague helping with the other children in the class. In most cases, (73%) the class was thus able to carry on with at least some of the normal programme of work, although 44% of classes also worked on tasks specially assigned for the SAT period. Nevertheless virtually all the teachers we questioned reported that the SATs caused major disruption to normal classroom practice and at least half felt that the SATs were totally unmanageable.

For the SATs alone the average time required was estimated at between 82 and 90 hours for planning, collecting resources, doing the assessments, marking and recording the results. In addition work for the rest of the class had to be prepared and planned as usual, in consultation with support staff, and many class teachers felt it necessary to go through children's work produced in their absence. In addition,

teachers reported that working intensively with small groups, constantly probing, questioning and recording, had proved unexpectedly exhausting.

Support arrangements were often in themselves a source of anxiety; Year 2 teachers were uncomfortably aware of the effect they had on the rest of the school. A teaching head in a rural school had given up all of her own non-contact time to release the Year 2 teacher; elsewhere all other teachers gave up their ancillary help. They willingly co-operated, but the knock-on effect was marked and some heads wondered how they would possibly manage if no major changes were made when Key stage 2 assessment was in process. Sometimes support teaching was - expensively - 'bought in' at an estimated cost of at least £500 in some cases, raising questions about the use of scarce resources in primary schools. Some teachers, too, were uncomfortable about the amount of supervision, if not actual teaching, undertaken by ancillaries, making such comments as, 'She's absolutely wonderful, but she's virtually in charge of most of the class for long periods and she's not paid to teach. It makes me feel that this is being done on the cheap.'

Different anxieties were expressed by teachers of horizontally and of vertically grouped classes. The first were weighed down by the sheer number of children who had to be assessed, but were at least able to feel that all children received an equal amount of attention. The second group constantly worried about whether the non-SAT children were being neglected, in spite of teachers' efforts to avoid this.

Teachers' anxiety also related to colleagues. 'I'm trying to be very careful in assigning levels to children,' said one teacher, 'but children perform so variably and they seem to forget so much during the summer holiday. If I say that Mary has achieved Level 3 for maths and next year she doesn't perform so well, how will her next teacher feel? We've always had such good relationships here; the last thing we want is friction over assessment.'

One more source of unmanageability for teachers was the requirement that they must avoid giving children direct help and 'asking questions that lead the child to the correct response.' They found this an unnatural way of working and felt that children could not understand the change in their practice.

Similarly, they usually encouraged co-operative, collaborative work in their classrooms and disliked having to reverse this practice. One teacher was heard to say, 'No, go away, Paul, I don't want you to see this.' Later she commented, 'He looked at me as if I'd gone mad!'

Teachers also worried about relationships with parents and the pressures they might be subject to. Several teachers reported that some parents had expressed anxiety, apparently envisaging paper and pencil tests with children passing or failing. While some parents were relieved to be reassured about this, others were shocked to see the quantity of materials schools had received and were concerned about the time SATs would require. Were children going to read to the teachers as often as usual? they asked: parents of Year 1 or Year 3 children in mixed classes were particularly anxious that their offspring might be neglected.

A problem arose in a school where several parents regularly helped in classrooms; a mother who did not do so emphasised that she objected to the presence of any parent in the room when her son was being 'tested', so all parental help for the Year 2 teacher was diverted to other classes.

Although parents were generally supportive, some made it clear that they saw National Curriculum assessments as a means for the government to check on teachers. In the words of one teacher, 'my teaching colleagues know what is involved in assessing children as Level 1, 2 or 3, but some parents in informal conversations have clearly perceived SATs as a test of the teachers. They've said, not unkindly, that they know we're all conscientious in this school, but that it will weed out some of those bad teachers.'

'All that work and hassle,' said one teacher, 'all that expense, all those resources that could have been poured into courses on assessment, and for what? So that children can be labelled as average, above or below average!'

In contrast, children appeared to have been protected from stress by the nature of the SATs and the efforts of their teachers to relate the assessment to good primary practice. Teachers were also clearly

anxious to make the tasks stress-free for children and drew on their professionalism and judgement to do so. On almost all occasions when SATs were observed, teachers introduced them by saying something like, 'Come and tell me about your game', or 'Let's go to the art room; we'll try an experiment.' Activities were presented as fun or as a part of normal classroom life; both observation and later interviews with children suggested that they were accepted as such.

Thus although 32% of teachers reported that some children showed signs of stress, most reported no special reaction (59%) and 63% of the teachers who responded suggested that the children experienced considerable enjoyment in doing the SATs.

Occasionally children demonstrated awareness of some element of assessment in the activities in such responses as 'Maybe she wanted to test us out on those things', or 'To see if we're good at it'. More often children happily admitted that they had no idea why they had been asked to carry out the activities: 'I don'[t know - she said "Four of you will be with me" - I don't know why'. They seemed to accept this, whether they described the tasks with enthusiasm - 'It was fun, I really liked it.' - or indifference - '... a bit boring, but quite nice'.

Teachers' strategies for avoiding stress in children included making no overt reference to, or demonstration of assessment in their presence. Children were not usually dismissed from a group if they were clearly finding the task too difficult, but were allowed to continue with teacher assistance which meant that they would not be considered to have reached the Level on which they worked. Sometimes they were told that they had worked extremely well, but that they looked tired and could 'leave this work for now'; the determination of most teachers to protect children from anxiety and from a sense of failure was impressive. One teacher even wondered, half-humorously, whether her children's relaxed attitude to the SATs was conducive to their producing the best work of which they were capable, or whether some awareness of being tested might have lent an edge to their performance.

In this respect it is important to contrast teachers' views of the teacher assessment component which most of them felt had caused little or no disruption to normal classroom routine (81%); was

reasonably accurate (97%) and manageable in terms of time (79%). At the same time these assessments had led to at least some new insights for many teachers (73%). Given that all the teachers felt there was a pretty good match between SATs and teacher assessment results, why bother to have SATs at all?

For the answer to this question we need to go back to the original blueprint for national assessment which was set out in the 1988 TGAT report. Here the rationale for national assessment was discussed in terms of four different purposes that national assessment should serve

- 'diagnostic' assessment to identify individual pupil's strengths and weaknesses;
- 'formative' assessment to give feedback and encouragement;
- 'summative' assessment to report on a given pupil's attainment at a given stage of schooling; and
- 'evaluative' assessment to provide aggregated information about the overall level of pupil achievement in any particular school as the basis for comparing one school with another.

It is immediately clear that SATs do not provide very well for the first two of these purposes. They are neither frequent enough nor sufficiently integrated into the normal routines and curricular emphases of the classroom to provide guidance for pupils and teachers about appropriate individual learning targets. The National Curriculum provides considerable flexibility as to the way in which it is to be implemented such that the particular skills and knowledge emphasised in a SAT may not match the pedagogical approach of the school. The fact that the SATs come only at the end of a key stage and that they identify very broad levels of attainment suggests that their real purpose is primarily summative and evaluative: to provide parents and the community as a whole with reliable information about children's achievements in a particular school. But do they in fact do this?

Although most teachers felt that the SATs were reasonably accurate (65%) only 3% of teachers felt they had learned a lot from the exercise. One respondent expressed the view of many: 'This is such a waste of time. I could be teaching now instead of doing all this to find out what I know already!'

Do teachers, then, trust SATs to measure children's attainment? In some cases teachers were concerned that they had underestimated children's abilities; elsewhere researchers talked to teachers who held the reverse view, that some pupils who met the criteria for reaching a level were being assessed over-generously, ('Well, he's managed all of these, so I must record him as Level 3, but he's not really a Level 3 child.') This clearly implies that some teachers may have a view of their pupils' achievement which is not affected by any evidence provided by SATs. This view may be based on an implicit form of norm-referencing with teachers using the more familiar categories of below average, average and above average children and linking these judgements to National Curriculum levels. There is clearly room for further investigation here. On the other hand, some teachers expressed pleasure at finding that children were achieving higher levels than they had anticipated - 'I didn't think she'd manage that.'

'Because all other teachers of Year 2 children will be using the same SAT, the results will be more standardised than your own continuous assessments' says the Handbook of Guidance for SAT (p.4). Responses from both teachers and headteachers in our survey schools identified a formidable list of reasons to make this assertion at best, questionable. The conditions in which children carried out the SATs varied so widely that it is difficult to see the results as reliable. Some groups worked in the corner of a busy and crowded classroom where their classmates frequently visited the group and made comments. In some classes disruptive children made disproportionate demands on the teacher's attention such that in some cases the SAT even had to be abandoned. Other children were able to work in a peaceful and otherwise empty room with the undivided attention of the teacher. Could it be confidently predicted that the groups would have worked in the same way if they had exchanged circumstances?

In other cases, the normal class teacher was absent and some children therefore had to be assessed for SATs by someone else which may well have affected their performance. Differences in the degree of LEA and school support provided for schools was also reported as a potential source of variability in

the ways the SATs were conducted, as was the difference in teacher approach caused in some cases by increasing confidence and familiarity on the part of the teacher and in others by teachers' flagging enthusiasm as they commenced the same task for the ninth time!

Such causes of real differences in pupil performance are more than matched by differences between teachers in the way in which they interpreted pupils' achievements in terms of the various levels. Indeed each level seems to include an extraordinarily broad range of attainment. In one classroom, for example, one group needed an hour to complete Sci 1 Part B. (The SAT is estimated to require about an hour if children move on to complete Part C, Level 3). This group was seen by their teacher as the most able; she stayed with them for about ten minutes, carefully explaining the first part of the activity and then moved about the classroom, returning to give help and reminders at each new phase - using the hand lens, weighing, etc. They worked with concentration and care, needing only brief interventions until they were individually questioned. None of them achieved two of the Level 3 criteria, however, so they did not move on to Part C and all were assessed at Level 2. The second group in the same class needed much more help and more detailed instructions; when the teacher moved away they all waited for her to return before making the next step and they constantly asked and told one another what to do. After forty minutes most of this group had completed only one line on the chart; they continued for well over another hour in the afternoon and by that time their teacher had decided to sit with them and to guide them through each new step. The Handbook of Guidance stresses that 'there is *no time limit* for an Activity. ... Some children and some groups *will* take longer.' The teacher felt that she had no alternative to assessing the second group also at Level 2, but she felt very uneasy about it. Certainly to the observer, the Level 2 assessment range seemed to encompass children of such widely different achievement as to be almost meaningless.

Another factor in the reliability of assessments which caused concern to teachers was the subjectivity involved in the judgements they were instructed to make. During the English 2, Part B SAT, for example, what were they to accept as 'a sensible prediction' about what might happen in the story child[illegible]ad been reading? How strictly should they interpret 'meaningful phrasing or intonation fluen[illegible]' One boy clearly understood and entered into the story he had read, relating it in

spontaneous conversation not only to his own experience but to imagined experience: what he would do if he had his own magic bunk bed. His teacher was worried, however, by his rather jerky and hesitant style of reading aloud: was this 'meaningful phrasing'? It seemed to be more closely connected to his natural speaking style than to his reading ability: should it play a part in this particular assessment? The boy was allowed to move on to Part D, where again fluency and 'appropriate expression' were required; children are also asked to 'read the story silently' and this proved almost impossible for this particular child. He kept breaking into speech: 'Hey, listen to this, it says she ...'. It was a delightful and lively response to the story but again his teacher was worried: would she be over-generous if she decided that he could, in fact, read silently on this evidence?

Although moderation has a key role to play in this respect it has proved difficult to implement across such a wide canvass of attainments, assessment purposes and aggregation rules. Our observations included the case of one girl whose knowledge and understanding of maths were felt by her teacher to be very good, but whose natural mode of work was slow and careful. Maths 3 Part B requires children to add and subtract to ten 'without any obvious counting or computation'. Her teacher was sure that the girl was not counting; she answered correctly but hesitantly. The moderator on this occasion insisted that answers should be given before a count of three and that this particular girl was too slow in responding to reach Level 2. What was being assessed here: mathematical ability, reaction times, speaking style or confidence? And what about the child who scores O for a particular attainment target because he was not assessed due to absence but whose aggregated level will nevertheless be affected by the inclusion of the O? Some teachers were also concerned as to whether moderation would apply to the way in which levels are interpreted across different key stages in order to provide continuity of standards.

One obvious answer to some of these problems is the kind of 'short sharp test' now envisaged by the Government for other key stages. But whilst such tests may superficially appear to meet the perceived need for 'objective' information about children's levels of achievement they will in practice have very limited utility and be likely to be quite counterproductive in terms of the other two avowed purposes of national assessment - the diagnostic and the formative. Paper and pencil tests can only measure at

best a rather limited range of skills. Can a written test really measure reading ability? (TES, 2.8.91) Certainly such tests will not be able to assess the full range of skills embodied in the Key Stage 1 curriculum such as oral work, practical work, problem-solving and planning.

Written tests would appear to offer relative ease of administration and greater comparability between schools. They are also likely to produce extremely limited information and to constrain the curriculum in a way that in terms of the Government's own national curriculum would be unacceptable. The results of our research so far argue strongly for the alternative approach of providing teachers with SAT-type assessment materials that they can use as an appropriate and genuinely integrated part of their teaching. The many strongly-worded criticisms of SATs which have been voiced are almost exclusively concerned with their mode of implementation and use and not with the tasks themselves. These have been generally appreciated. So have the increased liaison between colleagues, the increased contact with parents and the provision of moderation arrangements which teachers see as being a result of the more formal obligation to engage in teacher assessment. Teachers' apparent willingness to embrace this obligation creatively and enthusiastically despite their experience of SATs may yet prove to be the latter's most enduring legacy. It certainly says something both about the indomitable spirit and about the strength of the professional commitment of the teachers we have studied.